RAW SEQUENCE LISTING   DATE: 12/13/2001
PATENT APPLICATION: US/09/908,943A   TIME: 13:45:04

Input Set : **A:\00281A.txt**
Output Set: **N:\CRF3\12132001\I908943A.raw**

ENTERED

```
 4 <110> APPLICANT: Yan, Riqiang
 5        Tomasselli, Alfredo G.
 6        Gurney, Mark E.
 7        Emmons, Thomas L.
 8        Bienkowski, Mike J.
 9        Heinrikson, Robert L.
11 <120> TITLE OF INVENTION: SUBSTRATES AND ASSAYS FOR BETA-SECRETASE ACTIVITY
13 <130> FILE REFERENCE: 29915/00281A.US1
15 <140> CURRENT APPLICATION NUMBER: 09/908,943A
16 <141> CURRENT FILING DATE: 2001-07-19
18 <150> PRIOR APPLICATION NUMBER: 60/219,795
19 <151> PRIOR FILING DATE: 2000-07-19
21 <160> NUMBER OF SEQ ID NOS: 197
23 <170> SOFTWARE: PatentIn Ver. 2.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 2070
27 <212> TYPE: DNA
28 <213> ORGANISM: Homo sapiens
30 <400> SEQUENCE: 1
31 atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60
32 ggcacccagc acggcatccg gctgcccctg cgcagcggcc tggggggcgc ccccctgggg 120
33 ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
34 gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
35 gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300
36 gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360
37 taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
38 ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480
39 gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540
40 gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600
41 ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 660
42 cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 720
43 gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780
44 gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840
45 tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900
46 gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960
47 ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 1020
48 ttcccagtca tctcactcta cctaatgggt gaggttacca ccagtcctt ccgcatcacc 1080
49 atccttccgc agcaatacct gcggccagtg aagatgtgg ccacgtccca agacgactgt 1140
50 tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1200
51 ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260
52 catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320
53 gaagactgtg gctacaacat tccacagaca gatgagtcaa ccctcatgac catagcctat 1380
54 gtcatggctg ccatctgcgc cctcttcatg ctgccactct gcctcatggt gtgtcagtgg 1440
55 cgctgcctcc gctgcctgcg ccagcagcat gatgactttg ctgatgacat ctccctgctg 1500
56 aagtgaggag gcccatgggc agaagataga gattcccctg accacacct ccgtggttca 1560
57 ctttggtcac aagtaggaga cacagatggc acctgtggcc agagcacctc aggaccctcc 1620
58 ccacccacca aatgcctctg ccttgatgga gaaggaaaag gctggcaagg tgggttccag 1680
```

```
59 ggactgtacc tgtaggaaac agaaaagaga agaaagaagc actctgctgg cgggaatact 1740
60 cttggtcacc tcaaatttaa gtcgggaaat tctgctgctt gaaacttcag ccctgaacct 1800
61 ttgtccacca ttcctttaaa ttctccaacc caaagtattc ttcttttctt agtttcagaa 1860
62 gtactggcat cacacgcagg ttaccttggc gtgtgtccct gtggtaccct ggcagagaag 1920
63 agaccaagct tgtttccctg ctggccaaag tcagtaggag aggatgcaca gtttgctatt 1980
64 tgctttagag acagggactg tataaacaag cctaacattg gtgcaaagat tgcctcttga 2040
65 attaaaaaaa aaaaaaaaaa aaaaaaaaaa                                  2070
67 <210> SEQ ID NO: 2
68 <211> LENGTH: 501
69 <212> TYPE: PRT
70 <213> ORGANISM: Homo sapiens
72 <400> SEQUENCE: 2
73 Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val
74   1               5                   10                  15
76 Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser
77               20                  25                  30
79 Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
80           35                  40                  45
82 Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
83       50                  55                  60
85 Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
86   65                  70                  75                  80
88 Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
89               85                  90                  95
91 Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
92               100                 105                 110
94 Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
95           115                 120                 125
97 Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
98       130                 135                 140
100 Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile
101 145                 150                 155                 160
103 Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
104               165                 170                 175
106 Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp
107           180                 185                 190
109 Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro
110           195                 200                 205
112 Asn Leu Phe Ser Leu His Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln
113       210                 215                 220
115 Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile
116 225                 230                 235                 240
118 Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg
119               245                 250                 255
121 Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln
122           260                 265                 270
124 Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val
125           275                 280                 285
127 Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/908,943A

DATE: 12/13/2001
TIME: 13:45:05

Input Set : **A:\00281A.txt**
Output Set: **N:\CRF3\12132001\I908943A.raw**

```
128     290                     295                   300
130 Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp
131 305                     310               315                   320
133 Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr
134                     325               330                   335
136 Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val
137                 340               345               350
139 Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg
140             355                   360               365
142 Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala
143     370                     375               380
145 Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu
146 385                     390               395                   400
148 Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala
149                     405               410                   415
151 Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu
152                 420               425               430
154 Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro
155         435                 440               445
157 Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala
158     450                     455               460
160 Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp
161 465                 470                   475                   480
163 Arg Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp
164                 485                   490                   495
166 Ile Ser Leu Leu Lys
167             500
170 <210> SEQ ID NO: 3
171 <211> LENGTH: 1977
172 <212> TYPE: DNA
173 <213> ORGANISM: Homo sapiens
175 <400> SEQUENCE: 3
176 atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60
177 ggcacccagc acggcatccg gctgcccctg cgcagcggcc tggggggcgc ccccctgggg 120
178 ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
179 gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
180 gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300
181 gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360
182 taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
183 ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480
184 gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540
185 gggctggcct atgctgagat tgccaggctt tgtggtgctg gcttccccct caaccagtct 600
186 gaagtgctgg cctctgtcgg agggagcatg atcattggag gtatcgacca ctcgctgtac 660
187 acagcagtc tctggtatac acccatccgg cgggagtggt attatgaggt gatcattgtg 720
188 cgggtggaga tcaatggaca ggatctgaaa atggactgca aggagtacaa ctatgacaag 780
189 agcattgtgg acagtggcac caccaacctt cgtttgccca gaaaagtgtt tgaagctgca 840
190 gtcaaatcca tcaaggcagc ctcctccacg gagaagttcc ctgatggttt ctggctagga 900
191 gagcagctgg tgtgctggca agcaggcacc accccttgga acattttccc agtcatctca 960
192 ctctacctaa tgggtgaggt taccaaccag tccttccgca tcaccatcct tccgcagcaa 1020
```

```
193 tacctgcggc cagtggaaga tgtggccacg tcccaagacg actgttacaa gtttgccatc 1080
194 tcacagtcat ccacgggcac tgttatggga gctgttatca tggagggctt ctacgttgtc 1140
195 tttgatcggg cccgaaaacg aattggcttt gctgtcagcg cttgccatgt gcacgatgag 1200
196 ttcaggacgg cagcggtgga aggccctttt gtcaccttgg acatggaaga ctgtggctac 1260
197 aacattccac agacagatga gtcaaccctc atgaccatag cctatgtcat ggctgccatc 1320
198 tgcgccctct tcatgctgcc actctgcctc atggtgtgtc agtggcgctg cctccgctgc 1380
199 ctgcgccagc agcatgatga ctttgctgat gacatctccc tgctgaagtg aggaggccca 1440
200 tgggcagaag atagagattc ccctggacca cacctccgtg gttcactttg gtcacaagta 1500
201 ggagacacag atggcacctg tggccagagc acctcaggac cctccccacc caccaaatgc 1560
202 ctctgccttg atggagaagg aaaaggctgg caaggtgggt tccagggact gtacctgtag 1620
203 gaaacagaaa agagaagaaa gaagcactct gctggcggga atactcttgg tcacctcaaa 1680
204 tttaagtcgg gaaattctgc tgcttgaaac ttcagccctg aacctttgtc caccattcct 1740
205 ttaaattctc caacccaaag tattcttctt ttcttagttt cagaagtact ggcatcacac 1800
206 gcaggttacc ttggcgtgtg tccctgtggt accctggcag agaagagacc aagcttgttt 1860
207 ccctgctggc caaagtcagt aggagaggat gcacagtttg ctatttgctt tagagacagg 1920
208 gactgtataa acaagcctaa cattggtgca aagattgcct cttgaaaaaa aaaaaaa     1977
210 <210> SEQ ID NO: 4
211 <211> LENGTH: 476
212 <212> TYPE: PRT
213 <213> ORGANISM: Homo sapiens
215 <400> SEQUENCE: 4
216 Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val
217   1               5                  10                  15
219 Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser
220              20                  25                  30
222 Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
223          35                  40                  45
225 Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
226          50                  55                  60
228 Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
229  65                  70                  75                  80
231 Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
232              85                  90                  95
234 Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
235             100                 105                 110
237 Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
238             115                 120                 125
240 Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
241         130                 135                 140
243 Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile
244 145                 150                 155                 160
246 Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
247             165                 170                 175
249 Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Leu Cys Gly
250             180                 185                 190
252 Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly
253             195                 200                 205
255 Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu
256         210                 215                 220
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/908,943A

DATE: 12/13/2001
TIME: 13:45:05

Input Set : A:\00281A.txt
Output Set: N:\CRF3\12132001\I908943A.raw

```
258 Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val
259 225                 230                 235                 240
261 Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr
262                 245                 250                 255
264 Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu
265                 260                 265                 270
267 Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser
268                 275                 280                 285
270 Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val
271         290                 295                 300
273 Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser
274 305                 310                 315                 320
276 Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile
277                 325                 330                 335
280 Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln
281                 340                 345                 350
283 Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val
284                 355                 360                 365
286 Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala
287         370                 375                 380
289 Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu
290 385                 390                 395                 400
292 Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu
293                 405                 410                 415
295 Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser Thr Leu Met Thr
296                 420                 425                 430
298 Ile Ala Tyr Val Met Ala Ala Ile Cys Ala Leu Phe Met Leu Pro Leu
299         435                 440                 445
301 Cys Leu Met Val Cys Gln Trp Arg Cys Leu Arg Cys Leu Arg Gln Gln
302         450                 455                 460
304 His Asp Asp Phe Ala Asp Asp Ile Ser Leu Leu Lys
305 465                 470                 475
308 <210> SEQ ID NO: 5
309 <211> LENGTH: 14
310 <212> TYPE: PRT
311 <213> ORGANISM: Artificial Sequence
313 <220> FEATURE:
314 <223> OTHER INFORMATION: Description of Artificial Sequence: synthetic
315       peptide sequence
317 <400> SEQUENCE: 5
318 Lys Val Glu Ala Asn Tyr Glu Val Glu Gly Glu Arg Lys Lys
319   1               5                   10
322 <210> SEQ ID NO: 6
323 <211> LENGTH: 15
324 <212> TYPE: PRT
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <223> OTHER INFORMATION: Description of Artificial Sequence: synthetic
329       peptide sequence
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/908,943A

DATE: 12/13/2001
TIME: 13:45:06

Input Set : **A:\00281A.txt**
Output Set: **N:\CRF3\12132001\I908943A.raw**

```
L:435  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:473  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:497  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:521  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:545  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:592  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:692  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:695  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:728  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28
L:731  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28
L:925  M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41
L:1042 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49
L:1061 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50
L:1080 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51
L:1099 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52
L:1118 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:53
L:1137 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54
L:1156 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55
L:1175 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56
L:1194 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57
L:1213 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58
L:1232 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59
L:1251 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60
L:1270 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61
L:1289 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:62
L:1308 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:63
L:1327 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:64
L:1346 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65
L:1365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66
L:1384 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:1403 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68
L:1423 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:69
L:1442 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70
L:1461 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71
L:1480 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72
L:1499 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73
L:1518 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74
L:1537 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75
L:1556 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76
L:1575 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:77
L:1594 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78
L:1613 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79
L:1632 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80
L:1656 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:81
L:1680 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82
L:1704 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:1728 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84
L:1752 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85
```

VERIFICATION SUMMARY   DATE: 12/13/2001
PATENT APPLICATION: **US/09/908,943A**   TIME: 13:45:06

Input Set : **A:\00281A.txt**
Output Set: **N:\CRF3\12132001\I908943A.raw**

L:1776 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86
L:1795 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87